DOCUMENT RESUME

ED 458 800 FL 026 949

AUTHOR Griffee, Dale T.
TITLE Questionnaire Translation and Questionnaire Validation: Are They the Same?
PUB DATE 2001-02-00
NOTE 10p.; Paper presented at the Annual Meeting of the American Association for Applied Linguistics (St. Louis, MO, February 24-27, 2001).
PUB TYPE Opinion Papers (120) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Content Validity; *Questionnaires; Research Methodology; *Second Language Instruction; Second Language Learning; Surveys; *Translation

ABSTRACT

The purpose of this paper is to give evidence for the thesis that if teachers using a questionnaire as a data collection instrument have the questionnaire items translated from one language into another, they cannot assume that the translated items are valid simply because they were translated. Even if the original questionnaire items were validated, this does not change the situation because the validity is context specific and is not an abstract notion that transfers from instrument to another. A questionnaire written in one language and translated into another is not an equivalent survey instrument. Meaning and intention are part of what makes a questionnaire valid, and they are not as easily translated from language to another as the words are. The cultural context of the second language is different, and so the meaning and intention of the word will be understood differently to some degree. It is concluded that any questionnaire translated into another language must again be subjected to further analysis and pilot studies to confirm its validity anew. (Contains 16 references.) (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

Questionnaire translation and questionnaire validation: Are they the same?

Dale T. Griffee
South Plains College
<dalegriffee@AOL.com>

ED 458 800

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Dale T. Griffee

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL 026 949

BEST COPY AVAILABLE

Questionnaire translation and questionnaire validation: Are they the same?

Dale T. Griffee
South Plains College
<dalegriffee@AOL.com>

The purpose of this paper is to argue and give evidence for the thesis that if teachers using a questionnaire as a data collection method have the questionnaire items translated from one language (for example, English) into another language (for example, Japanese); they cannot assume that the translated items are valid simply because they were translated. Even if the original questionnaire items were validated, this does not change the situation because validity is context specific and is not an abstract notion that transfers from one instrument to another. I want to argue against the assumption that a questionnaire written in one language and then translated into another language results in an equivalent instrument. My point is not only must the original questionnaire items be validated, but the translated items must also be validated. This is an important issue because research is currently being done using questionnaire instruments created in one language (usually the language of the questionnaire designer) and administered in another language (usually the language of the persons for whom the questionnaire was designed).

Some researchers (Shimizu, 1995; Widdows & Voller, 1991) have made statistical inferences based on data from questionnaire items which were written originally in English, and then translated into Japanese. In both papers just cited, the purpose of the translated instruments was aimed at making inferences about students' needs or attitudes. Widdows and Voller wanted to know if their Japanese students' needs were being met by traditional teaching methods, and Shimizu wanted to know if her students had different attitudes toward Japanese teachers of English as opposed to foreign teachers of English. Neither of these studies gave reasons for the translation although we can suppose that they did so to ensure item comprehension by the Japanese participants in their studies, and thus increase validity. That translation of items from L1 to L2 enables comprehension by L2 native speakers is probably a widely held belief. For example, one colleague who read an earlier draft of this article pointed out that it seemed intuitively obvious that translation would be beneficial for some students.

Research such as that cited above raises the question of what are the minimum requirements for validation of questionnaires? I am claiming that for a questionnaire instrument, at least seven points must be accounted for in order to claim validation (Griffee, 2001). Validation for any data collection instrument is not concerned with achieving truth with a capital "T" or absolute objectivity. Nor is the honesty of the researcher being questioned. Although it is possible that some researchers have provided false information, I do not believe it happens on a regular basis. What does happen, though, is we researchers are often deluded. Our delusion

FL026949

is that we tend to believe that since we created the items on the questionnaire and understand the meaning and intention of those items, we assume that the persons taking our questionnaire also understand the meaning and intention of the items. To the extent that this is the case, then to that extent the items are valid, and the interpretation of the results is based on solid ground. But it is possible—indeed, probable—that some students have an understanding of certain items that is different from ours. It is this possibility that validation attempts to address. In the case of questionnaires, validation is the offering of evidence that the persons answering the questionnaire items understood what the items were asking in a way that is reasonably the same as the questionnaire developer. I will present these seven necessary conditions for validation in the form of seven questions.

1. Does the researcher state the purpose of the questionnaire? Purpose is necessary because we can't judge a questionnaire if we don't know why it is being used. In the body of an evaluation study or research paper it is typical to introduce the questionnaire with a statement along the lines of, "A questionnaire was given in order to ..." and then to finish the statement with the reason the questionnaire was given.

2. Is the construct defined? Unlike stating the purpose, a definition of the construct is almost never provided. By construct I mean the trait or attribute the questionnaire is designed to measure. As Most and Zeidner point out "before developing a test of any construct, one should clearly and explicitly express what one wants to test" (1995, p. 482). For example, if the questionnaire is designed to measure an attitude such as "liking", then liking is the construct. We need to know how is this attitude is being defined. Defining the construct is necessary for validation because when we administer a questionnaire, we use the results to draw conclusions e.g., the students liked the course or a certain aspect of the course. If we aren't explicitly clear what it means to like something, we are forced to rely on our implicit understanding which varies from person to person. Defining a construct can take more than one form. A strong form would be to offer a well documented theory which defines the construct. Many times, however, we can't find a theory. Another way of defining our construct would be to find a paper dealing with our subject in which the researcher defined the construct. In that case, we can borrow the definition and cite the paper. If all else fails, we can define the construct ourselves.

3. Are the items constructed in conformance with the construct? This means if the construct is defined as having qualities A, B, and C; we would like to know if there are items in the questionnaire which are designed to measure qualities A, B, and C (and not, say D). This is important because the questionnaire items need to have an established connection to the construct. In this way we can see that the items measure the construct, and not something else. For example, if our construct is "liking" and we define liking as having three components—an easiness factor, an enjoyment factor, and a learning promotion factor—we would expect to be told how

all the items related to one of those three subareas.

4. Was the questionnaire piloted, and were the results analyzed and reported? This is crucial because a pilot study provides the data for the evidence of validation. No pilot study means no data, and no data means no evidence to support the validation of the items. No data also means no indication of which items were not working well and should be eliminated or revised. Pilot data can be analyzed in many ways: Were any participants interviewed to see what they thought of the items, what kind of reliability evidence is provided (see Griffee, 1996 for a discussion of reliability), were any items correlated to confirm they were being answered in a consistent way, were the results factor analyzed? Finally, were any items eliminated or revised as a result of the pilot data?

5. Were descriptive statistics provided for the questionnaire results? After the questionnaire has been piloted and revised, it can be administered. After the administration, the test maker has to provide evidence showing how well the test is measuring the stated construct (Brown, 1988, p.101). In addition to a reliability coefficient, we would like to see descriptive statistics for the questionnaire as a whole as well as for each item. This is helpful because readers can see for themselves the overall results as well as how each item performed. If open ended items were included in the questionnaire and qualitative data were collected, we still need to see some form of data reduction, analysis and reporting of the qualitative data.

6. Is the questionnaire population defined? This means was information about the people who took the questionnaire provided? We might like to know the number of respondents, how they were selected, and enough description (e.g., age, gender, language proficiency, school affiliation, or other relevant categories) so that the reader can grasp who these people were who took the questionnaire. This is a necessary step because it goes to generalizability, or the ability of the reader to interpret the results and apply those results to his or her own situation.

7. Is the actual questionnaire provided? This is necessary so we can see and inspect the items, and also so the questionnaire can be used by other teachers. Not only does it not make sense to work this hard to create a questionnaire and then hide it away so that other teacher-evaluators can't use it, but providing the questionnaire allows replication and subsequent reanalysis which increases validation.

Here I wish to raise two questions, present evidence from the literature regarding the problems of translation, and finally offer some validation options. I will discuss the test instrument items from Shimizu (1995) and Widdows & Voller (1991). I will refer to the original English instrument as the E-doc and the resulting Japanese translation instrument as the J-doc. I will use the terms "test," "instrument," and "questionnaire" interchangeably.

Two questions The first question is, was the E-doc validated? In the two studies previously cited, the constructs underlying the questionnaire were not defined, no reliability information was provided, and no evidence was provided to show that the instruments were measuring the constructs. The answer is no. Therefore, the translation could not be valid since the original English language instrument on which it was based was not validated. Questionnaires with no validation are not unusual in the field of TESOL. For example, in Japan it is rare for articles reporting questionnaire data in *The Language Teacher* to report validation information. More recently, evaluators based in the UK (Karavas-Doukas, 1998; Kiely, 1998) used questionnaires with no mention of validation.

The second question is, was the J-doc validated? Remember, my argument is that the translated J-doc becomes a new instrument which has to be revalidated as a separate instrument. Even if the E-doc had been validated, the validation does not automatically carry over to the J-doc. We have to have additional evidence that the J-doc is measuring what it purports to be measuring. The mere fact that the students speak the Japanese language (ignoring the issue of foreign students in Japan) as their L1 and that the J-doc has been translated into the Japanese language is not enough. For example, if you are a native English speaker, you can ask yourself two questions. Have you ever read an English sentence which was translated from Japanese (or another language) into English and been left wondering what it was trying to say? The second question is, have you ever read an English sentence written originally in English by a writer you knew to be an English L1 writer, and still been left wondering what it meant? Most English native speakers can answer yes to both questions. This is why validation information is required in the first place. Even if the J-doc had been written originally in Japanese by a Japanese L1 writer, it would still require validation evidence.

Evidence from the literature Is it possible for questionnaire items to be validly translated? Many societies acknowledge, and in some cases, revere certain translated documents. Without translation, Christians would not have access to their scriptures, and the world would be without the understandings and wisdom supplied by classical thinkers such as Plato and Aristotle. In the modern era, bookstores regularly sell translated documents such as philosophical essays, novels, and poetry. While it is not my intention to call this practice into question, even here, things are not as obvious as they might first appear. Miller (1992) considers translation in the sense mentioned above, and discusses four problems encountered by virtually all translators: 1) the syntax of one language has no equivalent in another language, 2) words in one language don't have exact meanings in another language, 3) a word in one language has a spread of meanings that does not cover the spread of meanings in another language, and 4) words that can be used figuratively in one language cannot be used figuratively in another language. Miller concludes:

> Anyone who has translated will know the odd experience of being able to read and understand the original perfectly, as well as having native mastery of the target language, but of running constantly into unexpected and perhaps even insuperable difficulties in trying to turn the source text into the target language. The arrow keeps going awry and missing the target (1992, p. 124).

Suppose the E-doc were validated, what then? Widdows and Voller themselves suggest the difficulty if not impossibility of a valid translation. They state, "It is interesting to note that certain concepts quite fundamental to current EFL methodology proved impossible to render into straightforward Japanese" (1991, p. 128). They add that another difficulty arose from Japanese cultural understanding of learning styles. One item wanted to know if the student learned better when the teacher took an interest in them as a person. The problem was with the word "interest" because they found "it was impossible to eradicate entirely the connotation of sexual interest in the Japanese version" (Widdows & Voller, 1991, p. 128).

Evidence from the empirical studies In addition to the logic of validation, I turn to the results of two empirical studies which suggest that questionnaire items written in different languages carry different meanings, which raises serious problems with translation as a way of ensuring meaning. Yoshida (1990) conducted an experiment with second language learners (Japanese returnees who had lived in the U.S. for at least two years and had attended American schools). Thirty-five Japanese returnees were the experimental group, a group of 32 monolingual Japanese students in Japan were one control group, and a group of 21 monolingual American students in America were another control group. All three groups were given a word association task consisting of words from nature, daily life, society and ideas, and culture. The control groups answered in their own language. The experimental group was asked to respond in Japanese to the Japanese words and in English to the English words. The two lists of words were given in different order and a week apart. Yoshida compared the responses for each word, grouped the responses into semantic categories, and calculated the degree of agreement between the experimental group and each of the control groups. His analysis showed that for the society and ideas as well as the culture categories, "the bilingual group responded quite differently depending on which language they were using" (Yoshida, 1990, p. 22). For example, in giving word associations with the word "freedom" the experimental group gave responses such as responsibility, myself, human beings, and independence, words which did not appear at all with the Japanese translation.

In another study, Sakamoto (1996) investigated Hyland's (1994) use of translated questionnaire items adapted from Reid's (1987) learning style preferences questionnaire. Sakamoto's students were two groups of Japanese

women aged 20 to 22 years of age at Bunka Woman's University in Tokyo. Hyland had Reid's items translated from English to Japanese, and Sakamoto used these translated items except that she retranslated four of the items she thought misleading. Sakamoto administered both the English items and the translated items to her students allowing time between administrations to reduce the possibility of students simply remembering answers. She then compared the answers on the two questionnaires to see if the students answered the Japanese version differently than the English version and found that "about half of the 65 participants answered the same questionnaire statements differently in English and Japanese" (Sakamoto, 1996, p. 83). Sakamoto concludes:

> Clearly, then, there were differences between the questionnaire results in English and Japanese. The high discrepancy in this study warns us that the researcher should not simply consider translation as the answer to help the respondent understand the questionnaire better (1996, p. 87).

Validation options I am not arguing against translating questionnaire items although I have presented evidence that raises doubts that the items in L1 will be understood and answered in the same way as the items in L2. I am, however, strongly arguing that a translated questionnaire constitutes a different instrument which, in tern, has to be subject to its own validation. After piloting, analyzing, and revising items in the L2, a translated questionnaire can be administered and the results interpreted with a certain degree of confidence.

Another option would be to administer the questionnaire in L1 to students speaking a language other than the developer's L1. There are at least two cases in which this may a necessary option. One is in which the questionnaire is administered to respondents not all speaking the same language. This is a common situation in North American university classes. In this case it may not be practical to create and validate multiple questionnaires.

Another situation which might call for administering a questionnaire in a language not native to the respondents is when a visiting researcher goes to the country of the students. For example, a visiting researcher goes to Spain for a short time and does not have time to translate, pilot, and analyze items to create a Spanish language translation. In this case, there are certain validation strategies which still may be used to increase respondent comprehension. After questionnaire items have been created, ask a number of competent teachers to look at the items and make a judgment on their content. The results from this panel can be cited as validation evidence. In addition, show the questionnaire to a student panel, by which I mean a group of students similar to those for whom the questionnaire was developed. Be sure that you include lower level students on your panel. By student panel I do not mean that students must meet at the same time in the same place. In validating a questionnaire purporting to measure the construct "confidence" in speaking EFL (Griffee, 1997), I went to the school cafeteria and found some students

I recognized as having low English proficiency (former students). I showed them the questionnaire and asked them not to answer the questions, but to look at the items and circle any word they could not understand. A few items where checked, and I revised these items until subsequent student panels passed all items. Questionnaire validation involves more than a student panel, but showing items to a student panel is, I believe, a strategy which can be used to increase respondent understanding. Reporting the results from your student panel constitutes an additional kind of validation evidence.

I would like to conclude with two points. I am not trying to find fault with the two studies cited. I support Widdows and Voller as well as Shimizu in their research. Nevertheless, and this is my second point, we are now more informed as to what is involved in reliability and questionnaire validation. Among the things that we are currently aware of is that validation must be built into the design of the questionnaire from the very beginning (Luppescu & Day, 1990). We also know that piloting and analyzing data from the pilot must precede primary data collection. And we know that data resulting from questionnaires must also be analyzed and reported. To this list, I would add that we know that translation is not a short-cut validation solution. Translation results in a new document which itself must be piloted and analyzed.

References

Brown, J. D. (1988). Understanding research in second language learning. Cambridge: Cambridge University Press.

Griffee, D. T. (1996). Reliability and a learner style questionnaire. In G. van Troyer, S. Cornwell & H. Morikawa (Eds.), On JALT 95 Curriculum & Evaluation: Proceedings of the JALT 1995 International Conference on Language Teaching/Learning (pp. 283-292). Tokyo: The Japan Association for Language Teaching.

Griffee, D. T. (1997). Validating a questionnaire on confidence in speaking English as a foreign language. The JALT Journal, 19 (2), 177-197.

Griffee, D. T. (2001, January). Bare bones validation: A guide for teacher consumers of questionnaire research. Applied Linguistics Forum: Official newsletter of the TESOL Applied Linguistics Interest Section, 21 (1), 4, 11.

Hyland, K. (1994). The learning style of Japanese students. JALT Journal, 16, 55-74.

Karavas-Doukas, K. (1998). Evaluating the implementation of educational innovations: Lessons from the past. In P. Rea-Dickins & K. Germaine (Eds.), Managing Evaluation and Innovation in Language Teaching (pp. 25-50). Harlow: Addison-Wesley Longman.

Kiely, R. (1998). Programme evaluation by teachers: Issues of policy and practice. In P. Rea-Dickins & K. Germaine (Eds.), Managing Evaluation and Innovation in Language Teaching (pp. 78-104). Harlow: Addison-Wesley Longman.

Luppescu, S. & Day, R. (1990). Examining attitude in teachers and students: The need to evaluate questionnaire data. Second Language Research, 6 (2), 125-134.

Miller, J. H. (1992). Translation as the double production of texts. In C. Kramsch & S. McConnell-Ginet (Eds.), Text and context: Cross-Disciplinary perspectives on language study (pp. 124-134). Lexington, MA: D. C. Heath.

Most, R., & Zeidner, M. (1995). Constructing personality and intelligence instruments. In D. H. Saklofske & M. Zeidner (Eds.), International handbook of personality and intelligence (pp. 417-446). New York: Plenum Press

Reid, J. M. (1987). The learning style preferences of ESL students. TESOL Quarterly, 21, 87-109.

Sakamoto, M. (1996). The effect of translating survey questions. TUJ Working Papers in Applied Linguistics, 9, 82-88.

Shimizu, K. (1995). Japanese college student attitudes towards English teachers: A survey. The Language Teacher, 19 (10), 5-8.

Widdows, S., & Voller, P. (1991). PANSI: A survey of the ELT needs of Japanese university students. Cross Currents, 18 (2), 127-141.

Yoshida, K. (1990). Knowing vs behaving vs feeling. In L. A. Arena (Ed.), Language proficiency: Defining, teaching, and testing (pp 19-40). New York: Plannum Press.

Griffee, D. T. (2001, February). Questionnaire translation and questionnaire validation: Are they the same? Paper presented at the convention of the American Association of Applied Linguistics, St. Louis, MO. This is a revised version of a paper presented at the Japan Association of Language Teaching 1998 national conference in Omiya, Japan titled Translating questionnaires from English into Japanese: Is it valid? and published In Barfield, Betts, Cunningham, Dunn, Katsura, Kobayashi, Padded & Parry, (Eds.), On JALT'98 Focus on the classroom: Interpretations: Proceedings of the JALT 24th International Conference on Language Teaching/Learning & Educational Materials EAP. Tokyo: The Japan Association for Language Teaching.